

Copyright (c) 1993 - 2000 Compugen Ltd.									
GenCore version 4.5									
Run on: March 1, 2001, 16:16:37 ; Search time 91.75 Seconds									
(without alignments)									
15.839 Million cell updates/sec									
title: Perfect score: US-09-331-631A-3_COPY_29_73	Scoring table: BLOSUM62	Sequence: Gapop 10.0 , Gapext 0.5	searched: 88757 seqs, 32294092 residues	total number of hits satisfying chosen parameters: 88757	Minimum DB seq length: 0	Maximum DB seq length: 200000000	Post-processing: Minimum Match 0%	Maximum Match 100%	Listing first 45 summaries
database : SwissProt_39;*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
result No.	Score	Query Length	DB ID	Description	RESULT 1	STANDARD;	PRT;	588 AA.	ALIGNMENTS
1	77.5	30.8	588 1	VELB_GOSHI	VCLB_GOSHI	STANDARD;	PRT;	588 AA.	
2	72	28.6	605 1	VCLA_GOSHI	P09801;				
3	67.5	26.8	47 1	ACRP_LUCY	DT 01-MAR-1989 (Rel. 10, Created)				
4	67	26.6	1170 1	TSP1_HUMAN	DT 01-MAR-1989 (Rel. 10, Last sequence update)				
5	67	26.6	1170 1	TSP1_MOUSE	DT 15-JUL-1999 (Rel. 38, Last annotation update)				
6	65	25.8	1173 1	TSP1_XENLA	DE VICILIN_C72_PRECURSOR (ALPHA-GLOBULIN B).				
7	63	25.0	524 1	SBP_SOYBN	OS Gossypium hirsutum (Upland cotton).				
8	57	22.6	724 1	AT1_VACCV	OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.				
9	57	22.6	726 1	AT1_CAMFC	OC [1]				
10	57	22.6	1284 1	AT1_COMPX	RP SEQUENCE FROM N.A.				
11	55	21.8	305 1	PPP6_RAT	RA Chilan C.A., Pyle J.B., Legocki A.B., Dure L. III;				
12	55	21.8	795 1	ENPL_CHICK	RT Developmental biochemistry of cotyledon storage proteins and germination. XVII. DNA and amino acid sequences of the members of the storage protein families.;				
13	54.5	21.6	244 1	Y153_CAEEL	RT				
14	54.5	21.6	1021 1	Y188_CAEEL	RT				
15	54	21.4	291 1	CGD1_BRAKE	RT				
16	54	21.4	643 1	RQ60_CAEEL	RL FUNCTION: SEED STORAGE PROTEIN.				
17	53.5	21.2	301 1	G197_HUMAN	CC -!- SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN				
18	53.5	21.1	1170 1	TSP2_BOVIN	CC BODIES				
19	53	21.0	702 1	AT1L_VARV	CC -!- SIMILARITY: TO OTHER 7'S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONVICILIN, CONSYCININ, ETC.).				
20	52.5	20.8	91 1	UCRH_HUMAN	CC CONVICILIN				
21	52.5	20.8	1172 1	TSP2_HUMAN	CC VICILIN				
22	52	20.6	33 1	MRP1_MAIZE	CC				
23	52	20.6	37 1	CG2S_LUPIN	CC				
24	52	20.6	82 1	C2_OXYPN	CC				
25	52	20.6	716 1	ENPL_RABIT	CC				
26	52	20.6	802 1	ENPL_MOUSE	CC				
27	52	20.6	803 1	ENPL_HUMAN	CC				
28	52	20.6	804 1	ENPL_CANFIA	CC				
29	52	20.6	188 1	TIR_ECOLI	CC				
30	51.5	20.4	106 1	COLA_HORSE	CC				
31	51.5	20.4	108 1	COLB_HORSE	CC				
32	51.5	20.4	147 1	LAPP_HAEF	CC				
33	51.5	20.4	298 1	HISL_CANFIA	CC				

AC P09799;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE VICTILIN GC72-A PRECURSOR (ALPHA-GLOBULIN A).  
 OS Gossypium hirsutum (upland cotton).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Malvales; Malvaceae; Gossypium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chian C.A., Borroto K., Kamalay J.A., Dure L. III;  
 RT "Developmental biochemistry of cottonseed embryogenesis and alpha  
 germination. XIX. Sequences and genomic organization of the alpha  
 globulin (vicilin) genes of cottonseed.";  
 RL Plant Mol. Biol. 9:533-546(1987).  
 CC !- FUNCTION: SEED STORAGE PROTEIN.  
 CC !- SUBCELLULAR LOCATION: COMPLENDARY MEMBRANE-BOUNDED VACUOLAR PROTEIN  
 CC BODIES.  
 CC !- SIMILARITY: TO OTHER 75 SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,  
 CC CONVICILIN, CONGLICININ, ETC.).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M19378; AAA31069.1; --.  
 DR PIR; S06398; S06398.  
 DR ISSN; 15447; 1CAK.  
 DR INTERPRO; IPR00111; --.  
 DR PFAM; PF00546; Seeds; 7s; 1.  
 KW Seed storage protein; Signal.  
 FT SIGNAL 1 23  
 FT CHATN 24 605 AA; 71049 MW; C9DB9371C976953B CR64;  
 SQ SEQUENCE 605 AA;  
 Query Match 28 6%; Score 72; DB 1; length 605;  
 Best local Similarity 42.5%; Pred. No. 0.26;  
 Matches 17; Conservative 9; Mismatches 10; Indels 4; Gaps 3;  
 QY 1 SEFD-RQEEBECKRQCMQLETSGMRQCVSQCDKRFEEDEI 39  
 Db 33 SEDPQQQRVEDCRKRC-QLETRGTEQ - DKCEHRSSETOL 69  
 RESULT 3  
 AGRP\_LUFCA STANDARD: PRT; 47 AA.  
 ID AGRP\_LUFCA STANDARD: PRT; 47 AA.  
 AC P56568;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DE 6.5 KDA-ARGININE/GLUTAMINE-RICH POLYPEPTIDE (6.5K-AGRP).  
 OS Luffa cylindrica (Smooth Loofah) (Sponge gourd).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Cucurbitales; Cucurbitaceae; Luffa.  
 RN [1]  
 RP SEQUENCE.  
 RP TISSUE=SIEED;  
 RX MEDLINE=97357433; PubMed=9214759;  
 RA Kimura M., Park S.S., Sakai R., Yamashiki N., Funatsu G.;  
 RT "Primary structure of 6.5k-arginine/glutamate-rich polypeptide from  
 the seeds of sponge gourd (Luffa cylindrica).",  
 RL Biosci. Biotechnol. Biochem. 61:984-988(1997).  
 CC !- FUNCTION: STORAGE PROTEIN WHICH PROVIDES NITROGEN AND CARBON  
 CC RESERVES DURING GERMINATION AND SEEDLING GROWTH.  
 CC !- MASS SPECTROMETRY: MW=5093.39; METHOD=MALDI.

RESULT 4  
 TSPL\_HUMAN STANDARD: PRT; 1170 AA.  
 ID TSPL\_HUMAN STANDARD: PRT; 1170 AA.  
 AC P07996;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE THROMBOSPDIN I PRECURSOR.  
 GN THBS1 OR TSPI OR TSP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=ENDOTHELIAL CELLS;  
 RP MEDLINE=87057617; PubMed=2430973;  
 RX Lawler J., Hynes R.O.;  
 RT "The structure of human thrombospondin, an adhesive glycoprotein with  
 multiple calcium-binding sites and homologies with several different  
 proteins.", J. Cell Biol. 103:1635-1648(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88139590; PubMed=2918029;  
 RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,  
 RA Baumgartel D.M., Rotwein P., Frazier W.A.;  
 RT "Complete thrombospondin mRNA sequence includes potential regulatory  
 sites in the 3', untranslated region.", J. Cell Biol. 108:729-736(1989).  
 RN [3]  
 RP SEQUENCE OF 1-397 FROM N.A.  
 RX MEDLINE=87157592; PubMed=3030396;  
 RA Kobayashi S., Eben-Mocutchan F., Framson P., Bernstein P.;  
 RT "Partial amino acid sequence of human thrombospondin as determined by  
 analysis of cDNA clones: homology to malarial circumsporozoite  
 proteins.", R. Biochemistry 25:8418-8425(1986).  
 RN [4]  
 RP SEQUENCE OF 1-374 FROM N.A.  
 RX MEDLINE=86287276; PubMed=3461443;  
 RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;  
 RT "Characterization of a cDNA encoding the heparin and collagen binding  
 domains of human thrombospondin.", Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).  
 RN [5]  
 RP SEQUENCE OF 1-166 FROM N.A.  
 RX MEDLINE=9291870; PubMed=2545457;  
 RA LaFerriere C.D., German T.M., Dixit V.M.;  
 RT "Characterization of the promoter region of the human thrombospondin  
 gene DNA sequences within the first intron increase transcription.",  
 J. Biol. Chem. 264:11222-11227(1989).  
 RN [6]  
 RP SEQUENCE OF 1028-1170 FROM N.A.  
 RA La Fleur M., Jobin C., Gauthier J., Kreis C.G.;  
 RL Submitted (XXI-1992) to the EMBL/GenBank/DDBJ databases.  
 CC !- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND  
 CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN,

CC LAMININ AND TYPE V COLLAGEN  
 CC -!- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.  
 CC -!- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 VWF DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS  
 (WHICH BIND CALCIUM).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M25631; AAA36741.1; -;  
 EMBL; X041665; CAA28370.1; -;  
 EMBL; X14787; CAA32899.1; -;  
 EMBL; J04835; AAA6178.1; -;  
 EMBL; M99425; AAB59566.1; -;  
 PIR; A05172; A03172;  
 PIR; A25812; A28155;  
 PIR; A26155; A28155;  
 PIR; A30140; A30140;  
 PIR; A34274; A34274;  
 PIR; P35555; IEMO.  
 MIM: 188060; -;  
 INTERPRO; IPR000561; -;  
 INTERPRO; IPR00884; -;  
 INTERPRO; IPR001007; -;  
 PROSITE; P00008; EGF; 2;  
 PFAM; PF00090; tsp\_1; 3.  
 PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 PROSITE; PS01186; EGF\_2; 1.  
 PROSITE; PS05002; VWF\_C; 1.  
 DR PROSITE; PS01208; VWF\_C; 1.  
 KW Glycoprotein; Cell adhesion; EGF-like domain; Signal.  
 FT CHAIN 1 170  
 FT SIGNAL 1 18  
 FT DOMAIN 19 232  
 FT DOMAIN 316 373  
 FT DOMAIN 379 548  
 FT DOMAIN 549 690  
 FT DOMAIN 723 950  
 FT DOMAIN 951 1170  
 FT REPEAT 379 430  
 FT REPEAT 435 491  
 FT REPEAT 492 548  
 FT DOMAIN 549 587  
 FT DOMAIN 588 645  
 FT DOMAIN 646 690  
 FT REPEAT 723 758  
 FT REPEAT 759 811  
 FT REPEAT 782 817  
 FT REPEAT 818 840  
 FT REPEAT 841 878  
 FT REPEAT 879 914  
 FT REPEAT 915 950  
 FT SITE 926  
 FT DISULFID 270 270  
 FT DISULFID 274 274  
 FT DISULFID 551 562  
 FT DISULFID 556 572  
 FT DISULFID 575 586  
 FT DISULFID 592 608  
 FT DISULFID 599 617  
 FT DISULFID 620 644  
 FT DISULFID 650 663  
 FT DISULFID 657 676

INTERPRO; IPR00008; EGF; 2.  
 PROSITE; PS00090; tsp\_1; 3.  
 PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 PROSITE; PS01186; EGF\_2; 1.  
 PROSITE; PS05002; VWF\_C; 1.  
 DR PROSITE; PS01208; VWF\_C; 1.  
 KW Glycoprotein; Cell adhesion; EGF-like domain; Signal.  
 FT CHAIN 1 170  
 FT SIGNAL 1 18  
 FT DOMAIN 19 232  
 FT DOMAIN 316 373  
 FT DOMAIN 379 548  
 FT DOMAIN 549 690  
 FT DOMAIN 723 950  
 FT DOMAIN 951 1170  
 FT REPEAT 379 430  
 FT REPEAT 435 491  
 FT REPEAT 492 548  
 FT DOMAIN 549 587  
 FT DOMAIN 588 645  
 FT DOMAIN 646 690  
 FT REPEAT 723 758  
 FT REPEAT 759 811  
 FT REPEAT 782 817  
 FT REPEAT 818 840  
 FT REPEAT 841 878  
 FT REPEAT 879 914  
 FT REPEAT 915 950  
 FT SITE 926  
 FT DISULFID 270 270  
 FT DISULFID 274 274  
 FT DISULFID 551 562  
 FT DISULFID 556 572  
 FT DISULFID 575 586  
 FT DISULFID 592 608  
 FT DISULFID 599 617  
 FT DISULFID 620 644  
 FT DISULFID 650 663  
 FT DISULFID 657 676

FT CARBOHYD 248  
 FT CARBOHYD 360  
 FT CARBOHYD 708  
 FT CARBOHYD 1067  
 FT CONFLICT 84  
 FT CONFLICT 523  
 SQ SEQUENCE 1170 AA; 129412 MW; 69D3EDE5AE3A395E CRC64;

RESULT 5  
 "TSP1\_MOUSE  
 ID TSP1\_MOUSE  
 STANDARD; PRT; 1170 AA.  
 AC P35441;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1995 (Rel. 34, Last annotation update)  
 DE THROMBOSPONDIN 1 PRECURSOR.  
 GN THBS1 OR TSP1.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathii; Murinae; Mus.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92147683; PubMed=1371115;  
 RA Lawler J., Duquette M., Ferro P., Copeland N. G., Gilbert D. J., Jenkins N. A.;  
 RA "Characterization of the murine thrombospondin gene.";  
 RA Dixit V. M.;  
 RA "Characterization of mouse thrombospondin 2 sequence and expression during cell growth and development.";  
 RL Laherri C. D., O'Rourke K., Wolf F. W., Katz R., Seldin M. F.,  
 RN [3]  
 RP SEQUENCE OF 1-490 FROM N.A.  
 RX MEDLINE=9037556; PubMed=2298070;  
 RA Bornstein P., Alfifi D., Devaramaylu S., Framson P., Li P.;  
 RA "Characterization of the mouse thrombospondin gene and evaluation of the role of the first intron in human gene expression.";  
 RL J. Biol. Chem. 265:16691-16698 (1990).  
 CC -!- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN, LAMININ AND TYPE V COLLAGEN.  
 CC -!- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.  
 CC -!- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 VWF DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS  
 (WHICH BIND CALCIUM).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M62470; AAA50511.1; -;  
 DR EMBL; M62451; AAA50511.1; JOINED.











CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).  
CC  
CC  
DR EMBL; X87581; CAA60885.1; -  
DR ZFIN; ZDB-GENE-980526-176; CYCD1.  
DR INTERPRO; IPR000553; -  
DR PFAM; PF0134; cyclin; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW cyclin; cell cycle; Cell division.  
SQ SEQUENCE 291 AA; 33067 MW; FA5274CB1B46D5EF CRC64;

Query Match 21.4%; Score 54; DB 1; Length 291;  
Best Local Similarity 38.5%; Pred. No. 16;  
Matches 15; Conservative 7; Mismatches 11; Indels 6; Gaps 2;

Qy 10 ECKRQCMQ----LENS-GQMRRCVSQCDKRFEDIDWS 42  
Db 242 DCLRSQEQTESLLESSLRQAOQHISNETKVEEDVDSL 280

Search completed: March 1, 2001, 16:16:38  
Job time: 418 sec

